

LETAL, FILS PRODIGE DE L'IA

Gaëlle Recourcé
Kwaga

Forum du GFII - le 5 décembre 2017 - *l'IA washing*

TALN ET IA



TALN
parler,
écrire,
comprendre
des textes



VISION
voir,
interpréter
des images



ROBOTIQUE
bouger,
attraper des
objets

**Systemes
Expert**
raisonner,
déduire,
prévoir

70 ANS DE TALN

I

Le TALN est créé comme une branche de l'IA

1950

II

Le TALN et la linguistique-informatique

1990

III

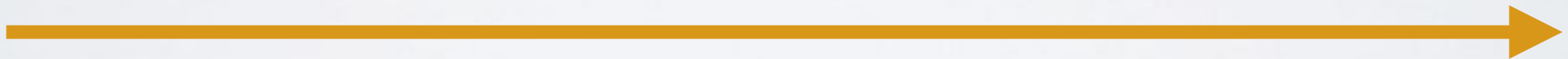
Le TALN s'ouvre au monde avec Internet

2000

Le TALN, un module de l'IA

2010

2014



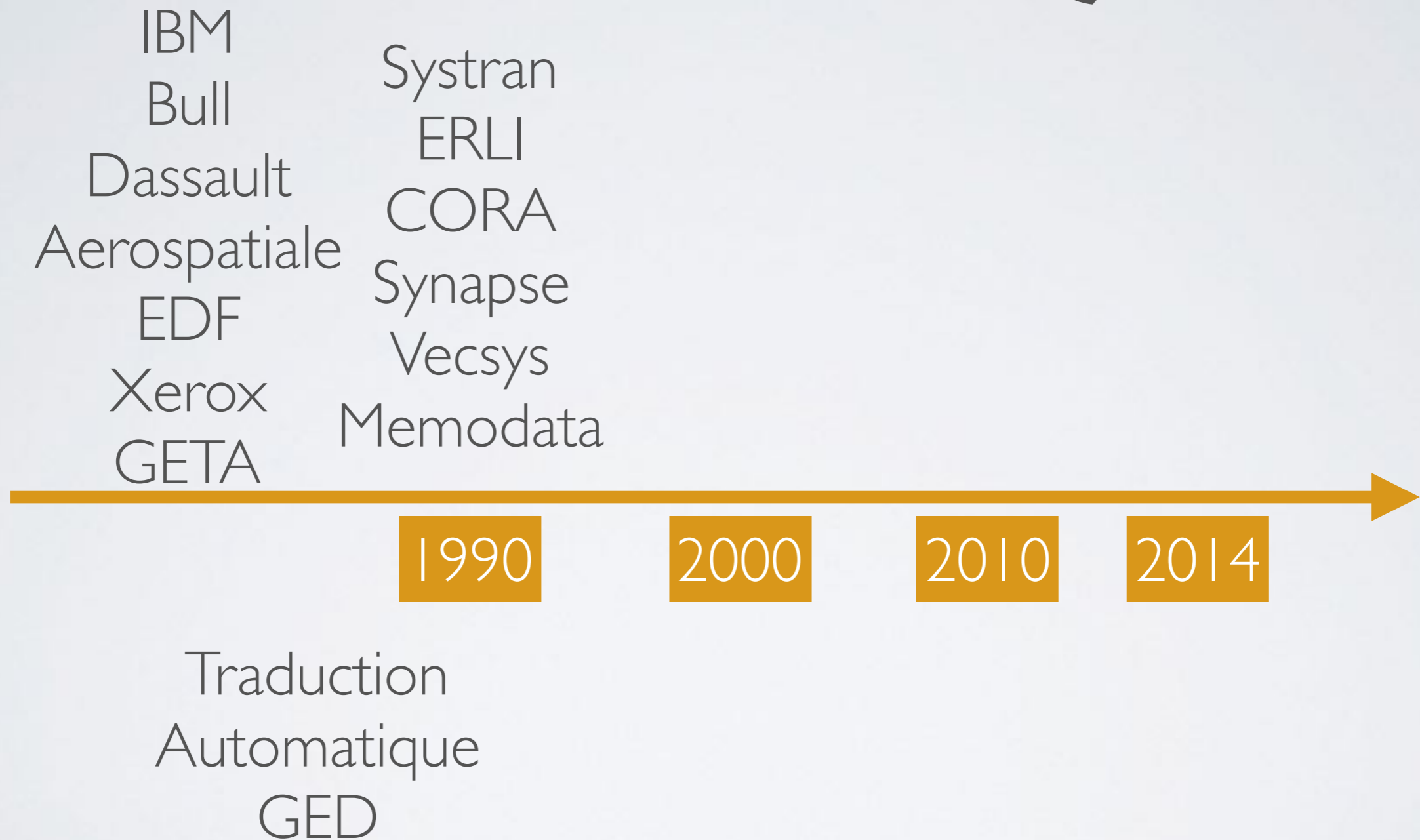
I - LES FONDATIONS DU TAL

- Se constitue autour d'une discipline nouvelle, la linguistique-informatique
- Définit un objet d'étude restreint, le texte numérique (ou la parole)
- Les données sont rares
- ... et les technologies limitées

ENJEU ORIGINEL - L'ANALYSE LINGUISTIQUE

- Des analyseurs linguistiques : phonétique, morphologie, syntaxe, sémantique mais **pas de pragmatique.**
- Une analyse « hors contexte » qui omet sciemment le locuteur, le lieu, le moment....
- Des applications pour les professionnels de la langue : correction orthographique, extraction de terminologie, construction de thesaurus, traduction automatique...

LES ENTREPRISES DE LA PREMIÈRE ÉPOQUE



II - LE TAL AU SERVICE DES DONNÉES MASSIVES

- Les années 90 et le développement de l'Internet
 - les données deviennent accessibles
 - on cesse d'imiter l'humain
 - les applications phare deviennent : moteur de recherche et fouille de données

LES ENTREPRISES DE LA SECONDE ÉPOQUE



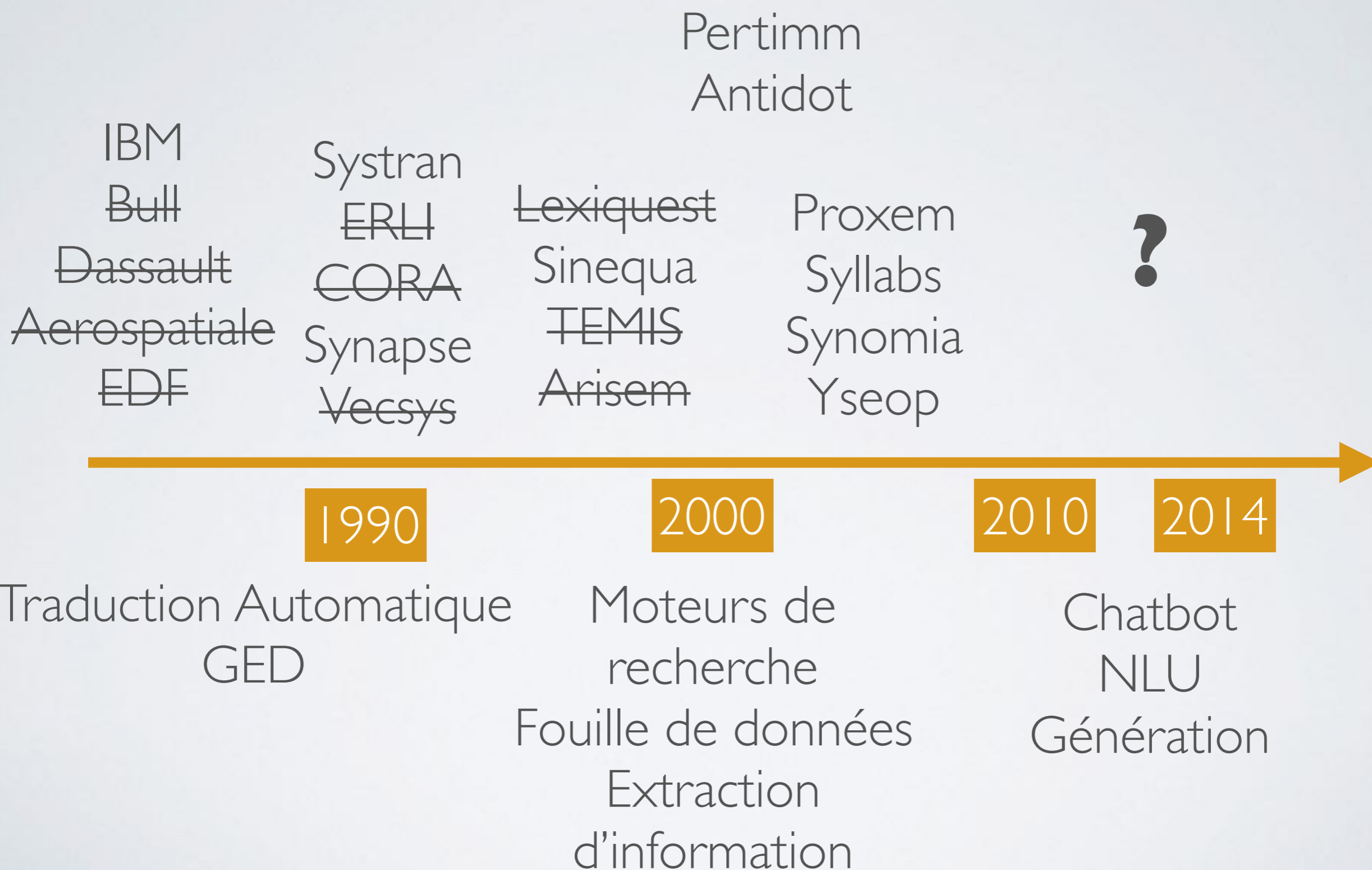
Traduction Automatique
GED

Moteurs de recherche
Fouille de données
Extraction d'information

III - LES RETROUVAILLES DUTAL ET DE L'IA

- Les « data » au coeur de ces retrouvailles
 - ➔ *Big data, Data science, Data scientists*
- Du NLP au NLU
- 2014-2015 - mise à dispo en ligne de services de *deep learning* (Google, Amazon, Facebook, IBM)

LES ENTREPRISES DE TAL-IA



Agents

Personal jam snips Wiidii	Professional Birdly Julie Desk Ullol ottspott	Platforms / Tools XPRATN clustaar. davi uella Botfuel Smartly.ai RECAST.AI
--	--	--

Robotics (Hardware + Software)

Robocortex
 IMMERSIVE ROBOTICS
 INTUITIVE ROBOTS
 BLUE FROG
 AXYN
 ALDEBARAN
 STARLET ROBOTICS
 EVENT BOTS
 HUMANOID
 STIMULACTIV MWOO
 octopus robots
 GOBOTO ROBOT
 Partnering ROBOTICS
 naïo

Enterprise

Security / Fraud QOSMOS Shift Technology CybelAngel ITrust Cyphame adiprom	Marketing / Sales tinyclues CustomerMatrix BEYABLE GEOMAIL Search xpr tastehit mediago	Customer support glogias CallDesk SENDERS viavoo DoYou DreamUp dictaNOVA	Business + Predictive Intelligence linkfluence PredicSis clustaar. entidot ZELROS DATA PRED IDMOG	HR Cluetree riminder
---	--	---	---	-----------------------------------

Platforms

Hardware / Processing Chronocam LightOn KALRAY CORTEX	Big Data / Machine Learning data iku SINEQUA tellmeplus DATA PRED craft ai antidot Invenis exenSa	Natural Language proxem syJLabs yseop	Vision / Image Chronocam Smartme up deepomatic heuritech	Audio niland ava OtoSense
--	--	---	---	---

Industries

AdTech DataBerries adomik	Insurance Shift Technology alan DreamQuark	Healthcare / Pharma Cardiologs HONESTICA Khresterion BIOSERENITY MENSIA implicity PIXYL DEXSTR Gynapse	Retail ALKEMICS verteego Moodstocks madd.ie tinyclues Neosensys deepomatic placemeter AntVoice Occi tastehit Search xpr BEYABLE
Legal Doctrine.fr	Manufacturing / Logietice BALYO energy efficiency VEKIA CORTEX		
Transportation NAVYA EASY MILE vulog karos Autokab	City + Infrastructures plume OtoSense placemeter DCbrain		

Agents

Personal 	Professional 	Platforms / Tools
-------------------------	-----------------------------	----------------------------------

Robotics (Hardware + Software)

Enterprise

Security / Fraud 	Marketing / Sales 	Customer support 	Business + Predictive Intelligence 	HR
---------------------------------	--------------------------------------	---------------------------------	---	---------------

Platforms

Hardware / Processing 	Big Data / Machine Learning 	Natural Language 	Vision / Image 	Audio
----------------------------------	--	-----------------------------	---------------------------	------------------

Industries

AdTech 	Insurance 	Healthcare / Pharma 	Retail
Legal 	Manufacturing / Logietice 		
Transportation 	City + Infrastructures 		

EXPERT EN TAL, POUR QUOI FAIRE ?

- pour pré-traiter les données textuelles (il serait bon de traiter aussi les autres données)
- pour structurer l'information (la connaissance, le comportement, modéliser le raisonnement)
- pour concevoir les chaînes de traitements des données : maintenabilité, scalabilité, thésaurisation des connaissances, feedback loop, multilinguisme

LIMITES, ENJEUX

- Les frontières entre les disciplines historiques (traitement de la langue écrite, de la parole, de l'image, de la vidéo, statistiques) persistent.
- Un flou sur les objectifs : de l'IA pour quoi faire ? Des raisonnements ? Une représentation des connaissances ? Des prédictions ?
- Question de la disponibilité des données utiles et des contraintes légales.

Merci de votre attention !